

EXHIBIT 7

Marc Egnal

Evolution of the Novel in the United States

The Statistical Evidence

This article examines the evolution of the novel in the United States using a remarkable new source, the Ngram database. This database, which spans several centuries, draws on the 15 million books that Google has scanned. It allows researchers to look at year-to-year fluctuations in the use of particular words. Using one of the available filters, the article is based on English-language books published in the United States between 1800 and 2008. But making sense of these data requires a framework. That framework is provided by the four periods that emerge from much recent writing on the novel. Four epochs—the sentimental era (1789–1860), the genteel era (1860–1915), the modern era (1915–60), and the postmodern era (1960–)—define the evolution of the novel and, more broadly, changes in American society and values. The article argues that a study of key words drawn from the Ngram database confirms the existence of these periods and deepens our understanding of them.

Ever since the first critical works appeared in the early US Republic, observers have proposed a variety of approaches to understanding the evolution of American literature. New points of view, new novels, and new bodies of evidence have led scholars repeatedly to refine or completely recast their interpretations. A database drawn from the more than 15 million books that Google has scanned offers just such an opportunity. This article argues that this source of information, called the Ngram database, clarifies and deepens the analysis of US novels. But databases never speak for themselves. They must be examined in the context of interpretive frameworks. This article

Social Science History 37:2 (Summer 2013)

DOI 10.1215/01455532-2074429

© 2013 by Social Science History Association

begins by exploring recent overviews of American literature. The article next discusses the nature of the Ngram database. It then suggests how these data shed light on American fiction from 1800 to the present.

The New Synthesis of American Literature

Quietly and without much fanfare, a new synthesis has reshaped analysis of American literature. This overview, which replaces older interpretations, provides a crucial background for understanding the Ngram data.¹ At first glance, the assertion that there is a current synthesis seems questionable. The emphasis in recent overviews has been on uncertainty, conflicting explanations, and multiple viewpoints. In his introduction to the seven-volume *Cambridge History of American Literature*, the editor, Sacvan Bercovitch (1994), underscores the demise of older patterns of analysis. “For this generation of critics and scholars,” he writes, “American literary history is no longer the history of a certain, agreed-upon group of American masterworks. Nor is it any longer based upon a certain, agreed-upon historical perspective on American writing.” Bercovitch (*ibid.*: 2) continues, pointing to the diversity of recent viewpoints, “In our times, in short, the study of American literary history defines itself in the plural, through volatile focal points of a multifaceted scholarly, critical, and pedagogic enterprise.” The single-volume, multiauthored *Columbia History of the American Novel* (Elliott 1991) echoes these cautionary themes. Recent discussions of the “canon”—the works that all scholars should read and teach—also highlight diversity, conflict, and the questioning of older authorities (*ibid.*: ix–xviii; Alberti 1995).

Yet all is not chaos. What most recent overviews have in common is an emphasis on periods, and not just any set but rather four particular epochs that define the evolution of US literature. In the introduction to the *Columbia History of the American Novel*, the editor, Emory Elliott (1991: xiv), notes: “There is clearly a chronological progression in the book with four historically organized sections introduced by a specialist in each period.” The four are “Beginnings to Mid-Nineteenth Century,” “The Late Nineteenth Century,” “The Early Twentieth Century,” and “The Late Twentieth Century” (*ibid.*). With variations in name, these four periods define the material in the *Cambridge History of American Literature* as well. The first era is the least likely to bear a formal name. But the second has been called the genteel era, the third the modern era, and the most recent the postmodern era. The newly

published *Cambridge History of the American Novel* (Cassuto 2011) highlights the same four divisions. It begins with part 1, “Inventing the American Novel,” which takes the story up to the Civil War. Next comes “Realism, Protest, Accommodation,” which carries this overview to World War I. The third section, “Modernism and Beyond,” examines writing through the 1960s. The final section, “Contemporary Formations,” opens appropriately with the essay “Postmodern Novels.”

This fourfold division is more than a useful set of labels. It ties literature to the evolution of American society in profound ways. The four periods track changes in the United States, as the nation metamorphosed from an overwhelmingly rural society (in the first era) to one at least initially dominated by small towns (in the genteel era); then to one where vibrant cities set the tone (in the modern era); and further to a time characterized by megalopolises, sprawling entities often with decaying cores (in the postmodern era). Similarly, this periodization embodies far-reaching shifts in values, beginning with an initial concern for womanly virtue (in the first era); evolving into the elaborate codes that marked genteel America; continuing with a less rigid set of moral guidelines while gender and racial hierarchies remained (in the modern era); and concluding with a period in which hallowed assumptions about race, gender, and authority gradually crumbled (in the postmodern era). The four-period synthesis suggests no judgment about the canon, the value of straight or queer literature, colonial or postcolonial attitudes, or male or female roles. In that sense it is very much in keeping with the recent, more open sensibility. But it does assert that any reading of literature should be set securely in the context of these epochs and the evolution of US society.

This article reinforces the current synthesis. The material presented below makes clear that this fourfold division, with its deep roots in American society and culture, provides a valuable framework for literary analysis.

The Ngram Database

The Ngram database sheds light on those eras. But before any source, and particularly one as far-ranging as this compilation, can be used, there is a need to consider its nature, strengths, and weaknesses. In 2004 engineers at Google began making digital copies of books. It is a task of Brobdingnagian proportions. Since the invention of the printing press, about 129 million books have been published; Google has now digitized over 15 million of

them. The project encompasses, along with English books, works in seven other languages. Most of the publications have appeared since 1800. English-language authors, for example, published only 500,000 works between 1500 and 1800.

From the volumes scanned, covering the years from 1500 to 2008, Google selected 5.2 million books as the foundation of the Ngram database. The researchers chose this group partly based on the quality of the optical character recognition (OCR) results. They also selected works with the most complete metadata, including date, place, and language of publication. According to a tabulation undertaken in 2010, the largest portion (67 percent) of the more than 500 billion words in this compilation are in English, with smaller collections in French (8 percent), Spanish (8 percent), German (7 percent), Russian (7 percent), Chinese (2 percent), and Hebrew (< 1 percent). Italian was added to the database after these numbers were computed (Michel et al. 2010, 2011).

Anyone visiting the site (ngrams.googlelabs.com) to run English-language word searches finds a short and not entirely satisfactory list of possible filters. The one used in this article is American English. It limits the search to English-language books published in the United States. These works, however, can be of many sorts, including legal or technical documents, reprints of foreign books, sermons, political tracts, and American fiction. Each book is considered only once, so obscure texts are weighted as heavily as best sellers. There are some virtues to dealing with this large basket of works, but before that question is tackled, the other filters for English-language books should be noted. The database can also be searched for British English (works published in Great Britain), English Fiction (fictional works in English without regard to place of publication), and English and English One Million (both dealing with the broader selection of English-language books). Readers can also examine the smaller Ngram databases for books in seven other languages.

Word searches generate graphs, such as the 17 presented in this article, but what exactly is measured? The numbers on the left axis are percentages indicating for each year the proportion of all words in the scanned books that match the term selected. So if you search for *the*, a very common word, the results show that in most years about 5 percent of all words (Google calls them “unigrams”) are a match. More typical results are much lower. For example, the use of *sublime* peaks early in the nineteenth century at about

0.003 percent. So of every 100,000 words in books published in the 1810s, 3 of them are *sublime* (Cohen 2010; Google Books Ngram Viewer 2011; Hotz 2010).

Google allows users of the database to conduct a finer-grained analysis, but that sorting must be done manually. For any word queried, Google provides a list of the books that include the term during a particular time period. So for *sublime* in the American English database for the 10 years beginning in January 1843, some 254 books are listed. All are published in the United States, but roughly half are reprints of works first issued elsewhere, chiefly in Britain. Some are literary works, for example, the stories of Edgar Allan Poe and the novels of the Anglo-Irish novelist Maria Edgeworth. Many of the volumes are reviews from the United States, such as the *Knickerbocker Review* and the *United States Democratic Review*, or reprints of British publications, such as *Punch* and the *Edinburgh Review*. In sum, the Ngram American English database is a useful handbook for what Americans wrote, but it is a better guide for what middle-class Americans were reading.²

This article examines, with an emphasis on novels, the four periods that emerge from recent overviews of American literature. The analysis highlights the changing popularity of various words as revealed in the Ngram American English database. The graphs presented begin in 1800, although Google provides data for earlier years. But the paucity of books published in the United States before the nineteenth century, and particularly the lack of works written by Americans, make the data from the 1780s and 1790s (or earlier) less reliable. For the four periods, I have used the labels and dates that made the best sense to me. Some studies posit the same divisions, but others might move the termini 10 or even 15 years in one direction or another. Since the underlying graph is always presented, the precise dating is less important than the overall trend line.

Sentimental Era, 1789–1860

During the sentimental era the American middle class, which had first emerged as a self-conscious group in the mid-eighteenth century, came to dominate society and shape its values. Like the other epochs, the sentimental era was a period of great change, but its unity came from the celebration of emotions and a determination to encourage new “proper” and pious patterns of behavior. In contrast to the genteel era with its elaborate, settled codes, the

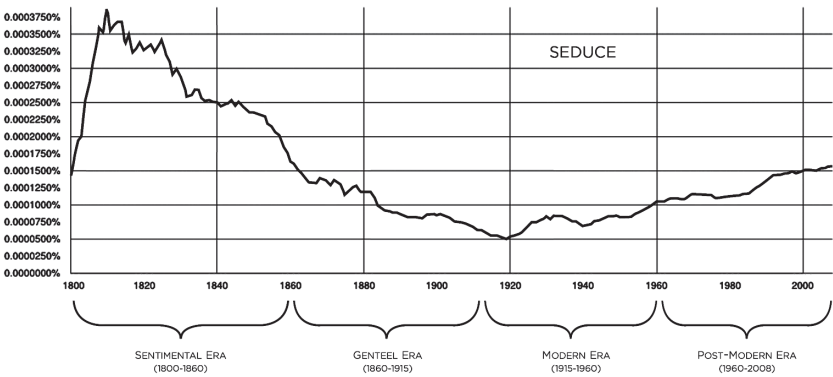


Figure 1 Frequency of *seduce* in American English database
Source: Ngram database.

emphasis in these years was on changing the way people acted (Brown 1940; Davidson 2004; Kete 2000; Tompkins 1985).³

The earliest American novels, such as William Hill Brown’s *Power of Sympathy* (1789) and Susanna Rowson’s *Charlotte Temple* (1791), are instructional texts for young women and broadcast the dangers of seduction, particularly by upper-class men. Charles Brockden Brown’s novels, such as *Ormond* (1799) and *Arthur Mervyn* (1800), similarly underscore these perils. While such warnings lessen after 1820, the horrors of seduction and the shame of the “fallen woman” remain evident in such works as Catharine Maria Sedgwick’s *Hope Leslie* (1827), James Fenimore Cooper’s *Deerslayer* (1841), and Nathaniel Hawthorne’s *Scarlet Letter* (1850). Not surprisingly, the literature of this era features the word SEDUCE more often than do later writings. (Capitalized words are displayed in the accompanying graphs.) The word FAITHFUL, a signifier for the proper sort of behavior, also recurs more prominently in this era than in later years (Bercovitch 1994: 620–60; Davidson 2004; Fiedler 1966: 105–23) (figures 1–2).

Efforts to check wanton misdeeds were the opening salvos in a campaign to reshape middle-class values. This concerted effort, which characterized the sentimental era, highlights the growing importance of Christianity and the cult of domesticity. Among the white population, church adherents (those who belonged to a congregation or often attended services) rose from about 14 percent in 1800 to 35 percent in 1850. Not surprisingly, CHURCH

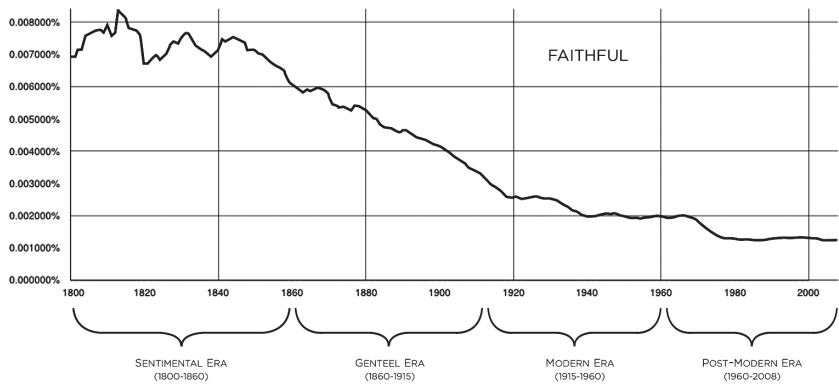


Figure 2 Frequency of *faithful* in American English database
Source: Ngram database.

was widely used during these years—far more so than in later epochs. Similarly characteristic of this first era was the emphasis given to the terms *PIOUS* and *SINFUL* (figures 3–4). The same holds true for the frequency of words (not here graphed) such as *Christianity*, *faith*, *God*, *grace*, and *prayer*. Certainly, the next period, the genteel era, lauded churchgoing—it was part of that epoch’s elaborate codes. However, the heightened concern for piety between 1789 and 1860 reflected the obsessions of a period of intense proselytizing. During these years ministers labored to change people’s beliefs. By contrast, in the genteel era most middle-class individuals knew their duties, and the flood of religious rhetoric receded.⁴

Ann Douglas (1988) describes the changes in this era as the “feminization of American culture” and emphasizes the role of ministers and women authors in this transformation (see also Baym 1993: 11–50; Finke and Stark 1986; Howe 2007: 285–327; Welter 1966). The growing importance of *MOTHER* in comparison to *FATHER* testifies to that feminization and reflects the ascent of the cult of domesticity (figure 5). In 1800 references to “father” are far more common than those to “mother,” but during the ensuing decades, and particularly after 1820, the gap closes strikingly.⁵

Key words highlight two other aspects of the sentimental era. One is the increased appreciation for the natural world. The newfound delight in nature was evident in the work of novelists; in the writings of transcendentalist philosophers, such as Ralph Waldo Emerson; and in the paintings of the Hudson River and Luminist schools. Much of the wonder that Americans felt in

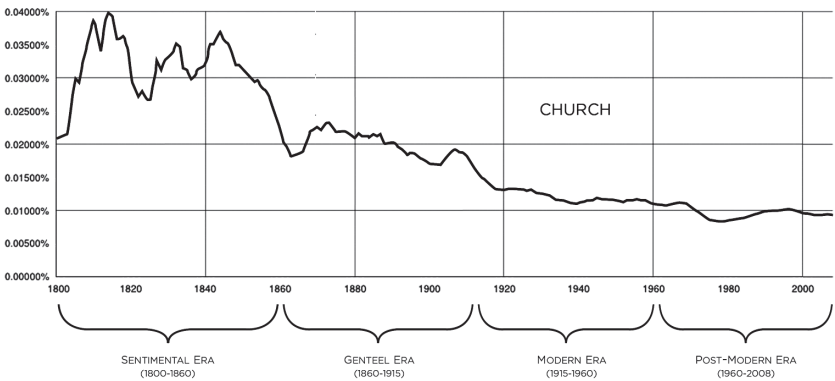


Figure 3 Frequency of *church* in American English database
Source: Ngram database.

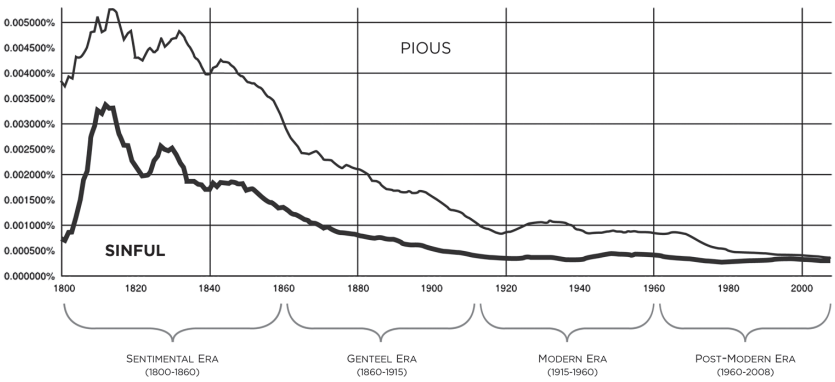


Figure 4 Frequency of *pious* and *sinful* in American English database
Source: Ngram database.

viewing their seemingly limitless forests, streams, waterfalls, and mountains is encapsulated in the word *SUBLIME* (figure 6). The term brings together the qualities of beauty with terror or awe. Many writers, such as Cooper, were fond of this concept. For example, *sublime* and its grammatical variants, such as *sublimity*, appear 20 times in *The Pathfinder* (1961 [1840]) and 8 times in *The Deerslayer* (1987 [1841]).⁶ After 1860 and particularly once industrialization and urbanization had darkened the American landscape, writers were less likely to view the natural world in that exalted light (Blakemore 1986; Novak 2007; Ringe 1971: 3–33).

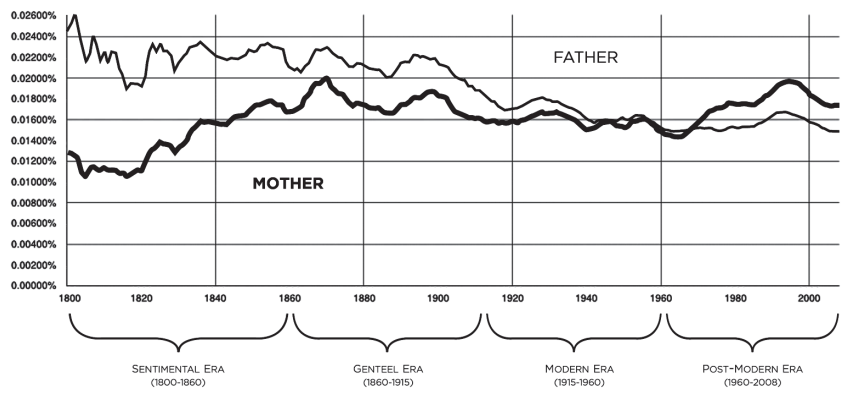


Figure 5 Frequency of *father* and *mother* in American English database
Source: Ngram database.

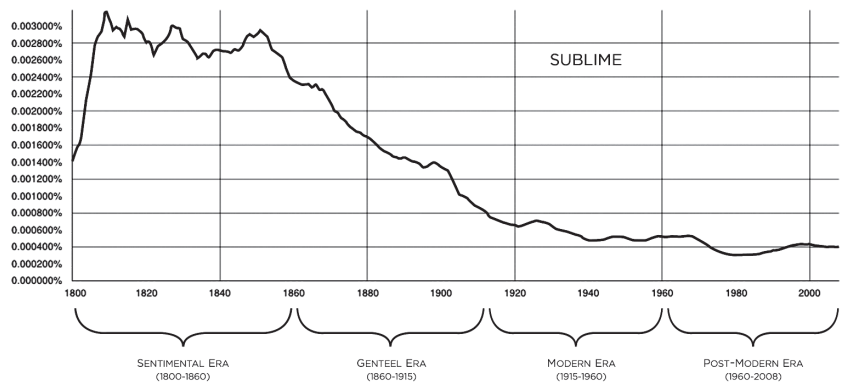


Figure 6 Frequency of *sublime* in American English database
Source: Ngram database.

Still another development important for both literature and, more broadly, American society was the rising importance of the concept of CHILDHOOD (figure 7). The colonists typically had regarded children as miniature adults. During the sentimental era middle-class families gradually came to view these early years as a special stage of development. Beginning in the 1830s, the new outlook was evident in the writings of transcendentalists such as Emerson, the pronouncements of ministers such as Horace Bushnell, and the works of such novelists as Hawthorne, Harriet Beecher Stowe, and Susan Warner. Childhood remained a serious time; not until the genteel era

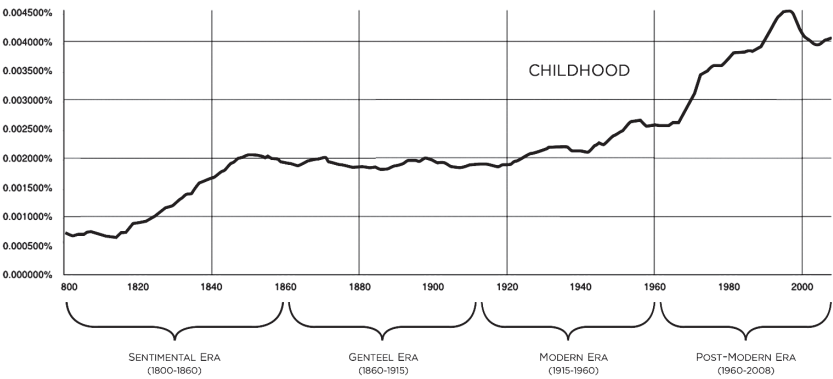


Figure 7 Frequency of *childhood* in American English database
Source: Ngram database.

were adults willing to give children periods of pure play. Still, a new appreciation of the young had emerged (and as the graph suggests, belief in childhood intensified in the postmodern era and will be discussed below) (Mintz 2004; Reinier 1996).

The Genteel Era, 1860–1915

The next epoch, the genteel era, was likewise marked by great changes. What characterized this period was the dominance of an elaborate if unwritten code that defined proper middle-class behavior. This worldview gradually broke down during the last decades of the era, but at its height this credo was marked by a long list of attributes. These included optimism and a clear-cut morality. The genteel code also embraced a belief in the superiority of the white race and a commitment to Protestantism, progress, the doctrine of the two spheres, temperance, and sexual modesty. In 1911, when George Santayana (1967 [1911]) wrote his celebrated essay “The Genteel Tradition in American Philosophy,” many had come to view these beliefs as oppressive. But in its heyday—the 1860s and the 1870s—this outlook encompassed many of small-town America’s noblest virtues, such as an emphasis on caring families and close-knit communities (ibid.; Smith 1978).

Perhaps no word characterizes the goodness of the early genteel era more than *LOVING*, a word widely used from the 1860s to the end of the century (figure 8). The term is the hallmark of Louisa May Alcott’s *Little Women*

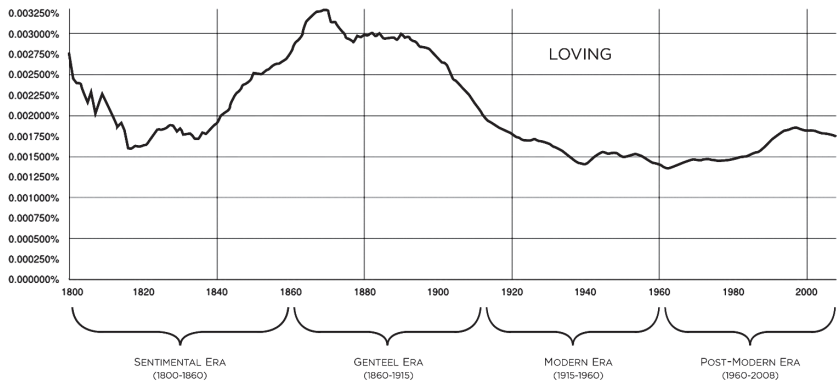


Figure 8 Frequency of *loving* in American English database
Source: Ngram database.

(2008 [1868]) and Mark Twain’s *Tom Sawyer* (1972 [1876]), two novels that epitomize these optimistic decades. Both works are set in small towns. While not everyone is well mannered in Alcott’s Concord or Twain’s Saint Petersburg, concern for family and neighbors lies at the heart of the books. *Loving* appears 44 times in *Little Women*. When Mr. March goes off to war, he writes to his wife, “I know they [his daughters] will remember all I said to them, that they will be loving children to you . . . that when I come back to them I may be fonder and prouder than ever of my little women.” On another occasion Mrs. March and the girls help a poor German family: “‘That’s loving our neighbor better than ourselves, and I like it,’ said Meg, as they set out their presents while their mother was upstairs collecting clothes for the poor Hummels” (Alcott 2001 [1868]: 52, 58). In *Tom Sawyer*, when the boys return from Jackson’s Island, Huck receives a warm welcome. The narrator comments, “The loving attentions Aunt Polly lavished upon him were the one thing capable of making him more uncomfortable than he was before” (Twain 1972 [1876]: 144).⁷

Other markers of genteel behavior come from the terms of reproach SCOUNDREL and COQUETTE (figure 9). Both denote individuals whose behavior was judged improper but who also remained accepted outliers in Victorian America. In Twain’s *Huckleberry Finn* (2003 [1884]) *scoundrel* is used repeatedly to refer to the two con men, the Duke and the Dauphin. The Duke in fact hurls the epithet at the Dauphin, while Huck refers to both grifters that way. Unlike *scoundrel*, whose use peaked in the genteel

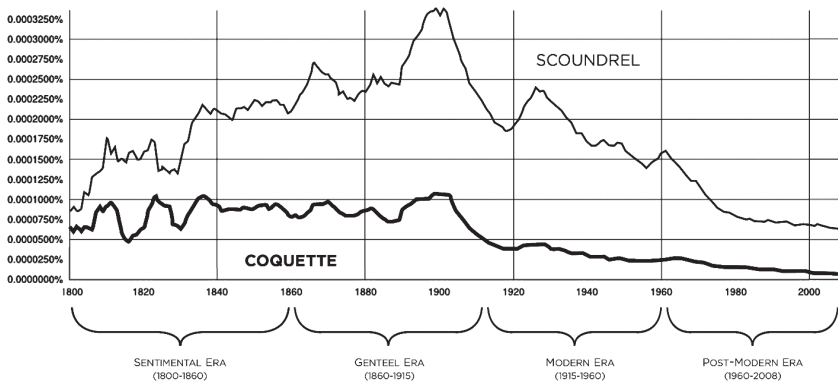


Figure 9 Frequency of *scoundrel* and *coquette* in American English database
Source: Ngram database.

era, *coquette* was widely used in the preceding era as well. Hannah Foster titled her 1797 novel *The Coquette*. Similar to *coquette* was *flirt*, a term whose use soared during the genteel era. In *Little Women* the word appears 17 times both as a noun and as a verb and usually hedged with disapproval. After 1900 the frequency of *scoundrel*, *coquette*, and *flirt* falls off sharply. The framework that cast into sharp relief these mildly transgressive behaviors was itself crumbling.

Another intriguing pair of words, LEG and LIMB, further illustrates the nature of the genteel era and the changes in this period. Until the 1880s the two terms tracked each other, with *limb* serving as a politer referent for *leg*. For example, in *Little Women* Alcott introduces both words. *Limb* is typically used when describing the March sisters. The novel notes Jo's "long limbs" and Beth's "feeble limbs" or how Amy sat "with every limb gracefully composed" (Alcott 2001 [1868]: 48, 241, 312). But when discussing little children, infants, or animals, *leg* becomes the preferred term. Alcott (ibid.: 365, 457) writes, for example, about the "baby's legs" and "Demi's short plaid legs." As the graph makes clear, references to limbs plummet after 1880 (figure 10). This decline, it would seem, is the first indication of the breakdown of the elaborate strictures that came together in the genteel era. More broadly, these changes (and others evident by 1900) suggest an arc that transcends and ties together the four epochs. From the 1780s until the late nineteenth century, American society was marked by an ever more well-defined social code. Beginning at the end of the nineteenth century and continu-

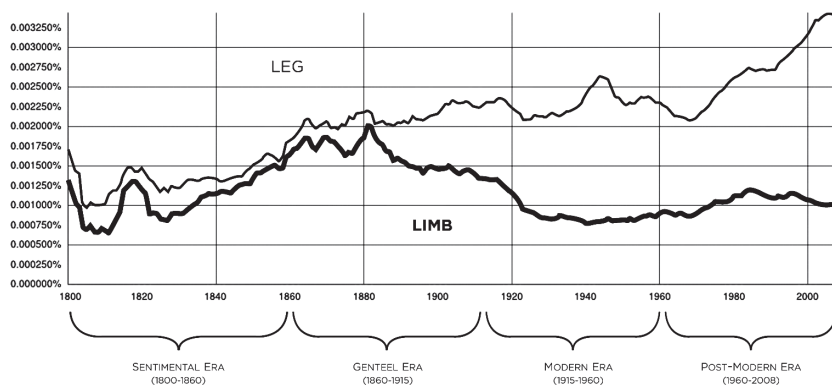


Figure 10 Frequency of *leg* and *limb* in American English database

Source: Ngram database.

ing through the modern and postmodern eras, these prescribed patterns of behavior gradually broke down and were replaced by guidelines that were increasingly less restrictive.⁸

By the turn of the century the genteel world was clearly crumbling, even as the Victorian codes continued to matter. The suicides of the female protagonists in Kate Chopin's *Awakening* (1899) and Edith Wharton's *House of Mirth* (1905) illustrate the force those ideas retained. Still, the decline of older values was unmistakable. Industrialization, urbanization, and the "new" immigration from eastern and southern Europe all undercut the hegemony of small-town, Anglo-Saxon America, which had provided the foundation for the era. A genteel America remained after 1915, lingering particularly in small towns and rural areas, but it no longer dominated the larger society (Delbanco 1997: 83–154; May 1959).

Modern Era, 1915–1960

The changes that sounded the death knell for the genteel era also marked the opening of the modern era. Cities, not small towns, now ruled the social landscape. Many in the middle class no longer deemed sinful those activities, such as drinking or sexual relations outside marriage, that earlier generations had condemned as immoral. Mainstream writers, such as Sinclair Lewis, now reviled rather than praised small towns. But while modern America was in many respects freer and more daring than genteel society had been, it

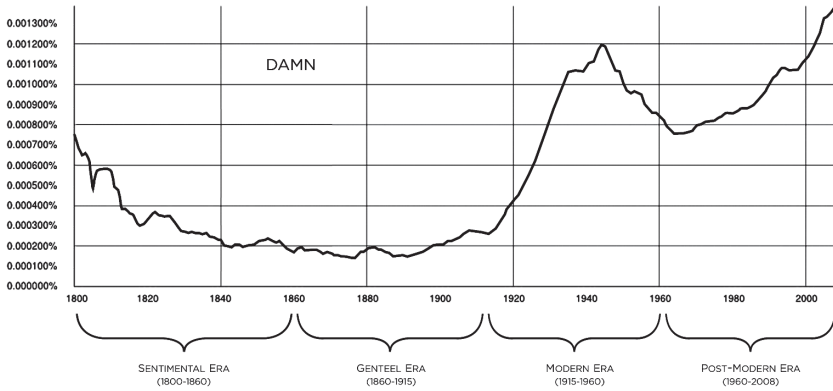


Figure 11 Frequency of *damn* in American English database

Source: Ngram database.

still bore the unmistakable stamp of an established order. Although women now voted and the Harlem Renaissance briefly flourished, gender and racial hierarchies held firm. Few questioned the value of families or religion, even while these institutions were redefined. Even during the Great Depression few doubted that the United States was a wonderful country or that the American dream was worth pursuing (Adams 1978; Baker 1987; Deen 2002; Douglas 1996; Kaplan 2010; Rado 1994).

In part, the modern era was defined by the new freedoms that emerged with the breakdown of genteel values. The decline of *coquette* and *flirt* signaled a new sexual openness. Those who now regarded the “loving” families of Victorian America as stifling institutions celebrated the possibilities that came from franker, more open relationships. The disappearance of *scoundrel* pointed to a society in which individuals undertook grimmer deeds. Few would label as “scoundrels” William Faulkner’s mean-minded protagonists, such as Popeye in *Sanctuary* (1931) or Joe Christmas in *Light in August* (1932). The villains of genteel fiction pale in comparison to the psychologically twisted characters in modern works (Allen 1952; Susman 1984).

The rising use of DAMN is another sign of the loosening of older values (figure 11). This imprecation may seem mild to today’s ears, but the expression, a shortening of *God damn*, was blasphemous and rarely used in literature or public forums before 1915. When Rhett Butler told Scarlett O’Hara in *Gone with the Wind* (1939), “Frankly, my dear, I don’t give a damn,” at least some viewers were shocked (Flamini 1975: 317–20). (The 1930 Motion Pic-

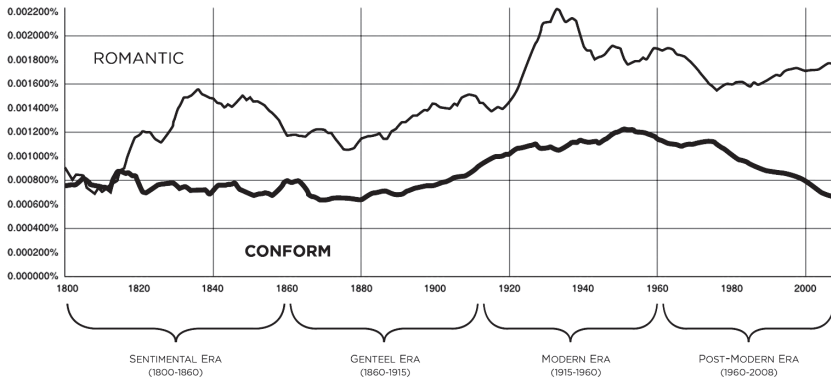


Figure 12 Frequency of *romantic* and *conform* in American English database
Source: Ngram database.

ture Association's Production Code had banned the word and fined the producer David O. Selznick \$5,000.) The word abounds in modern literature. For example, *damn* or *damned* appears 43 times in F. Scott Fitzgerald's *This Side of Paradise* (1920) and 35 times in his *Beautiful and the Damned* (1922).

The new sense of freedom that marked the early years of the modern era was also captured in the word **ROMANTIC** (figure 12). Where *loving* suggested the dependable, caring individuals who were the pillars of the Victorian family, *romantic* had a more transgressive feel. It described persons who listened to their own hearts and cast aside society's codes. *Romantic* had first risen in importance during the heyday of transcendentalism. Its frequency lessened during the early genteel era, when Victorian morality prevailed. But as those strictures weakened, *romantic* became more common, and its use soared in the 1920s. Jay Gatsby in Fitzgerald's *Great Gatsby* (1925) stands as the quintessential romantic hero. In introducing him, the narrator, Nick Carraway, notes that Gatsby possesses "a romantic readiness such as I have never found in any other person and which it is not likely I shall ever find again" (Fitzgerald 1925: 2).

The modern era had another side, one that was more staid and conformist. The ascent of the term **MIDDLE CLASS** also characterizes this period (figure 13). This designation had been gradually growing in importance since the nineteenth century, but the gains were most striking between 1915 and 1940 (another sharp advance came after 1960). The modern era was an epoch of increasing wealth and comfort despite the hardships of the Great Depres-

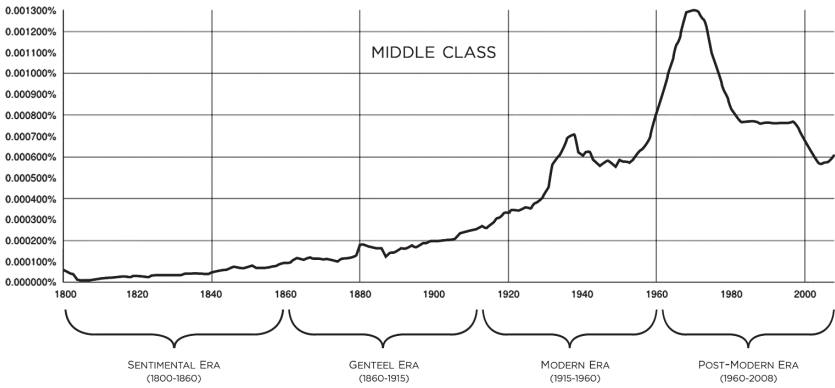


Figure 13 Frequency of *middle class* in American English database

Source: Ngram database.

sion. The Ngram database indicates that the use of words such as *profits*, *earnings*, and *investments* skyrocketed in this era. These terms, to be sure, were found more often in the business press than in fiction, but novels also help detail the expansion of the middle class. Logan Killicks and Jody Starks in Zora Neale Hurston’s *Their Eyes Were Watching God* (1937) are exemplars of these new monied individuals, as are the Snopeses in Faulkner’s trilogy that begins with *The Hamlet* (1940).

By the 1950s the middle class was prosperous, often smug, and prone to conformity. Significantly, citations of two transgressive terms—*romantic* and *damn*—which had risen in the 1920s and the 1930s, fell off in the 1950s. Instead, the last part of this epoch was marked by the steady ascent of the term CONFORM, whose use peaked in the 1950s (figure 12). In the 1950s novelists such as Ralph Ellison, J. D. Salinger, Mary McCarthy, and Saul Bellow depicted and criticized the middle class. But no writer saw an alternative to accepting and living with this oppressive society. For example, in Salinger’s *Catcher in the Rye* (1951: 13) the protagonist, Holden Caulfield, rails against conformist society, complaining that he was “surrounded by phonies,” but in the end plans to become part of that world.

Postmodern Era, 1960–

The postmodern era marks another sea change in American society and culture. Long-held assumptions that had shaped modern society now slowly

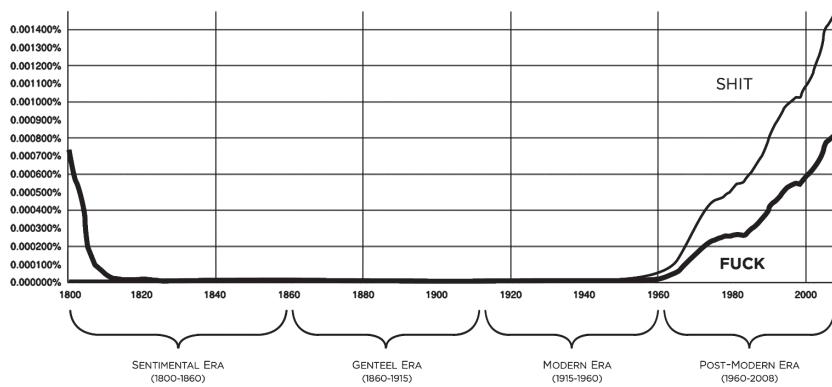


Figure 14 Frequency of *shit* and *fuck* in American English database

Source: Ngram database.

crumbled, much as the edifice of Victorian values had shattered early in the twentieth century. No longer did educated, middle-class Americans accept gender and racial hierarchies that had prevailed for centuries. The civil rights and women's movements had a far-reaching impact. (Terms such as *racism* and *sexism*, little used in earlier years, now soared in importance.)⁹ Novelists and others questioned the meaning and value of middle-class life. Characteristic of the era was the waning of the traditional family and the growing acceptance of alternative lifestyles. At the same time, the hollowing out of American industry eliminated a broad swath of the middle class. Changes in government policies heightened the extremes in wealth. Once vibrant cities now housed the gated rich and the teeming poor (Elias 1999; Jameson 1991, 1998; McHale 1987; Steiner 1999).

The growing disdain for older, middle-class values was evident in the spread of profanity. The prevalence of SHIT, FUCK, and other formerly taboo words testified to the new order (figure 14). When Norman Mailer wrote his first novel, *The Naked and the Dead* (1948), his publishers persuaded him to use *fug* instead of *fuck*. Mailer's postmodern works have no comparable circumspection. *Fuck* and *cocksucker* fill the pages of *The Executioner's Song* (1979). Similar language is present in other works of great seriousness, such as the novels of Philip Roth. (Incidentally, the spike between 1800 and 1815 in the graph for *fuck* reflects a flaw in Google's OCR: the older letter *s* was misread as *f*.)

Another change signaled by language is the growing importance of

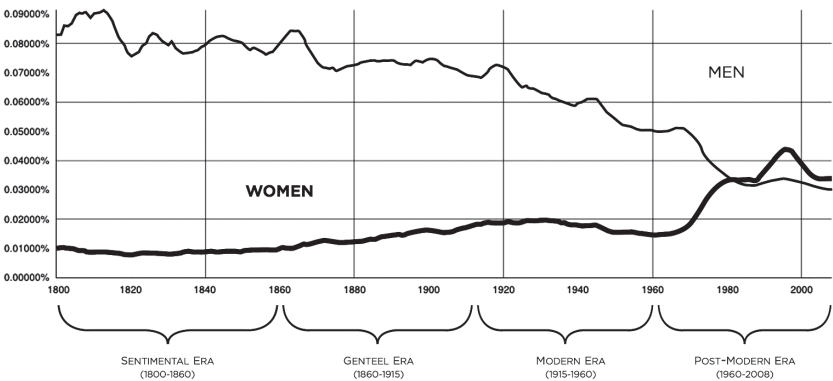


Figure 15 Frequency of *men* and *women* in American English database
Source: Ngram database.

women in society and literature. This trend is evident in the relative frequency of **WOMEN** and **MEN** in printed material (figure 15). During the sentimental era, even as *mother* grew more important, *women* barely rose in frequency. However, since 1960 *men* and *women* have switched places in the frequency of citations. The swap occurred because there were fewer references to “men”—a long-term trend—and because there were many more citations of “women.” *Man* or *men* no longer served as a synecdoche for *people*. Moreover, the use of *mother* has surpassed that of *father* in postmodern books.¹⁰

A similar pattern emerges in looking at *HE* and *I*, with *I* emerging as clearly predominant in the postmodern era (figure 16). References to the two pronouns had long fluctuated together, with *I* holding a slight edge since the 1830s, when transcendentalists and others exalted the individual. In today’s world *I* has become ever more common. The trend is hardly surprising in an era of confessional novels and, more recently, bloggers eager to post their innermost thoughts.

Finally, one of the most striking developments of the postmodern era has been the celebration of ties between parents and children, an emotional bonding reflected in many current novels. To many observers, postmodern society appears chaotic, lacking the structures and rules that once held groups together. In this sometimes frightening new world, the bedrock that remains for many is the love between a parent and a child. The new concern for children is evident in the sharply increased references after 1960 to

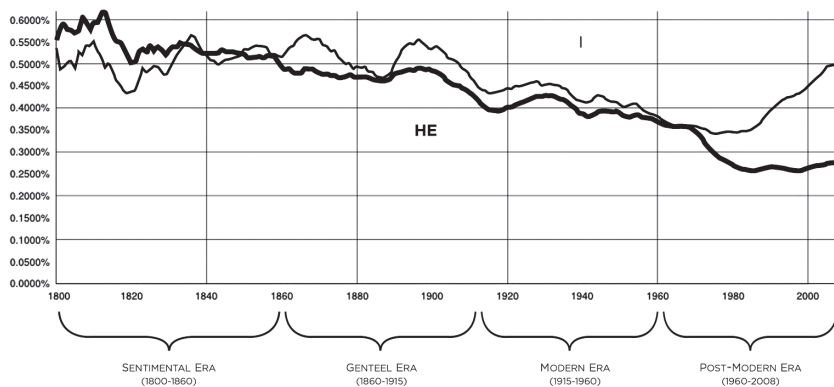


Figure 16 Frequency of *I* and *he* in American English database

Source: Ngram database.

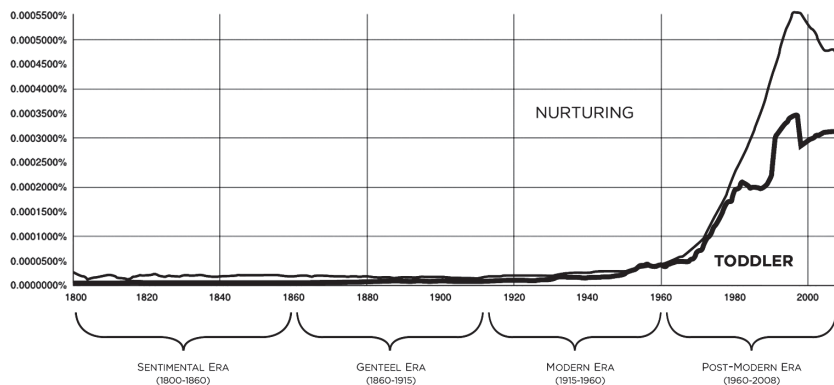


Figure 17 Frequency of *nurturing* and *toddler* in American English database

Source: Ngram database.

NURTURING and TODDLER (figure 17). The trend is also evident in the reinvigoration of the term *childhood*, which first rose in importance during the sentimental era. Many of the key postmodern novels have at their core the tight link between a parent and a child. Works that illustrate this tie are as diverse as Toni Morrison's *Beloved* (1987), Roth's *American Pastoral* (1997), Richard Russo's *Empire Falls* (2001), and Cormac McCarthy's *Road* (2006).¹¹

To conclude, this article argues that the study of various terms drawn from the Ngram database deepens our understanding of the evolution of American novels. Any essay such as this one, which deals with a broad sweep

of literature, a powerful but imperfect database, and links between social change and literature, can hardly be taken as the final word. The conclusions set forth are necessarily hypotheses subject to further testing. Future refinements to the Ngram database or its successors will surely make possible more precise searches.¹² Still these data, which rest on the analysis of millions of books, are too valuable to ignore even in their present form. These data confirm the current division of US literature into four eras and sharpen the conclusions about each of those periods.

Notes

I would like to thank Art Redding, Randolph Roth, Karen Wilson, and Brett Zimmerman for their comments on an earlier version of this article. Ben Egnal created the graphs.

- 1 Typically, earlier overviews sought to unify American literature around a single broad theme, such as the conflict between the “people” and plutocrats (Parrington 1927), the impact of mechanization (Marx 1964), or the differences between English and American novelists (Chase 1957; Fiedler 1966).
- 2 The Ngram database reflects the interests of the middle class more than the outlook of the working class. Google apparently scanned very few of the hundreds of thousands of dime novels that circulated in the nineteenth century and little of the pulp fiction or popular romance literature of the twentieth century. As various scholars have noted, these mass circulation works often present viewpoints at odds with “middlebrow” and “highbrow” literature. For a discussion of this popular literature, see Denning 1987; Goulart 1972; Jones 1978; Reynolds 1989.
- 3 The year 1789, which opens this era, marks the publication of the first American-authored novel, William Hill Brown’s *Power of Sympathy*.
- 4 Other words whose use peaked during this period and which reflect the religious fervor of the sentimental era are *Christ*, *confess*, *devil*, *Lord*, *repent*, and *Satan*. Analysis based on the *Oxford English Dictionary* (Simpson 2012) suggests that the meanings of these words, and indeed of most of the terms examined in the graphs, text, and notes of this article, have not notably changed since the 1780s. For the exceptions to the rule, see the analysis of *leg* and *limb* in the text and the discussions in notes 9 and 11. However, as this article makes clear, the *frequency* with which words have been used has varied dramatically (*ibid.*).
- 5 Other words characteristic of the sentimental era and the newly ascendant cult of domesticity are *affectionate*, *compassion*, *domestic*, and *submissive*.
- 6 E-book editions, available at www.gutenberg.org, made possible the tabulations of Cooper’s word use. This same source was used for the other single-book word counts mentioned in the text.
- 7 An extensive literature examines Twain’s complex relationship to the genteel tradition. See Brooks 1920; Kaplan 1966; Krauth 1999; and Smith 1978: 102–27.

- 8 Various terms whose use declines after 1900 are also markers of the breakdown of genteel values. These words include *brave*, *hero*, *honor*, *house*, *respectful*, *sentiment*, *sigh*, *tender*, and *villain*. *Temperance* also falls off but has a brief second surge in the 1920s with Prohibition.
- 9 According to the *Oxford English Dictionary* (Simpson 2012), *racism* with its modern meaning dates from the 1920s, while *sexism* as currently used emerged only in the 1960s. In earlier eras *sexism* was a rarely used term that referred to belonging to one gender, not discrimination against women (ibid.).
- 10 Similarly, *grandmother* has surpassed *grandfather* in the postmodern era.
- 11 The frequency of other terms that reflect the new devotion to nurturing has soared in the postmodern era. Among these words are *baby*, *babysitter*, *caregiver*, *caring*, *day care*, *infant*, and *nanny*. Two words, *domestic* and *home*, widely used in the sentimental and genteel eras, have experienced a new vogue. The *Oxford English Dictionary* (Simpson 2012) suggests that *babysitter* dates from the 1930s, while only during the nineteenth century did *toddler* (see figure 17) come to mean a small child.
- 12 Readers may wonder whether all the words tested for this article conformed neatly to the four periods. Most did, but a few did not. Some words have multiple meanings. For example, a *rake* can be a cad or a garden implement, and its Ngram pattern shows little clarity. Some words reflect events or particular time periods. *Soldier* peaked during the Civil War and the two world wars; *tenement* soared during the early twentieth century. A few words were simply puzzling. For example, *greedy* and *understanding* fluctuated with little evident pattern.

References

- Adams, Robert Martin (1978) "What was modernism?" *Hudson Review* 21 (1): 19–33.
- Alberti, John, ed. (1995) *The Canon in the Classroom: The Pedagogical Implications of Canon Revision in American Literature*. New York: Garland.
- Alcott, Louisa May (2001 [1868]) *Little Women*. Peterborough, ON: Broadview.
- Allen, Frederick Lewis (1952) *The Big Change: America Transforms Itself, 1900–1950*. New York: Harper.
- Baker, Houston A., Jr. (1987) *Modernism and the Harlem Renaissance*. Chicago: University of Chicago Press.
- Baym, Nina (1993) *Women's Fiction: A Guide to Novels by and about Women in America, 1820–1870*, 2nd ed. Urbana: University of Illinois Press.
- Bercovitch, Sacvan, ed. (1994) *The Cambridge History of American Literature*. Vol. 1, 1590–1820. New York: Cambridge University Press.
- Blakemore, Stephen (1986) "Language and world in *The Pathfinder*." *Modern Language Studies* 16 (3): 237–46.
- Brooks, Van Wyck (1920) *The Ordeal of Mark Twain*. New York: Dutton.
- Brown, Charles Brockden (1962 [1799]) *Ormond*. New York: Hafner.
- (1969 [1800]) *Arthur Mervyn*. Kent, OH: Kent State University Press.

- Brown, Herbert R. (1940) *The Sentimental Novel in America, 1789–1860*. Durham, NC: Duke University Press.
- Brown, William Hill (1969 [1789]) *Power of Sympathy*. Columbus: Ohio State University Press.
- Cassuto, Leonard, ed. (2011) *The Cambridge History of the American Novel*. New York: Cambridge University Press.
- Chase, Richard (1957) *The American Novel and Its Tradition*. New York: Doubleday Anchor.
- Chopin, Kate (1899) *The Awakening*. New York: Stone.
- Cohen, Patricia (2010) “In Five Hundred Billion Words, New Window on Culture.” *New York Times*, December 16, www.nytimes.com/2010/12/17/books/17words.html?page-wanted=1&_r=1.
- Cooper, James Fenimore (1961 [1840]) *The Pathfinder*. New York: New American Library.
- (1987 [1841]) *The Deerslayer*. New York: Penguin.
- Davidson, Cathy N. (2004) *Revolution and the Word: The Rise of the Novel in America*, expanded ed. New York: Oxford University Press.
- Deen, Stella, ed. (2002) *Challenging Modernism: New Readings in Literature and Culture, 1914–1945*. Burlington, VT: Ashgate.
- Delbanco, Andrew (1997) *Required Reading: Why Our American Classics Matter Now*. New York: Farrar, Straus and Giroux.
- Denning, Michael (1987) *Mechanic Accents: Dime Novels and Working-Class Culture in America*. London: Verso.
- Douglas, Ann (1988) *The Feminization of American Culture*. New York: Doubleday Anchor.
- (1996) *Terrible Honesty: Mongrel Manhattan in the 1920s*. New York: Noonday.
- Elias, Amy J. (1999) “And the beat goes on: Defining postmodernism and its other.” *Contemporary Literature* 40 (1): 161–80.
- Elliott, Emory, ed. (1991) *The Columbia History of the American Novel*. New York: Columbia University Press.
- Faulkner, William (1931) *Sanctuary*. New York: Cape and Smith.
- (1932) *Light in August*. New York: Smith and Haas.
- (1940) *The Hamlet*. New York: Random House.
- Fiedler, Leslie A. (1966) *Love and Death in the American Novel*. New York: Dell.
- Finke, Roger, and Rodney Stark (1986) “Turning pews into people: Estimating nineteenth century church membership.” *Journal for the Scientific Study of Religion* 25 (2): 180–92.
- Fitzgerald, F. Scott (1920) *This Side of Paradise*. New York: Scribner.
- (1922) *The Beautiful and the Damned*. New York: Scribner.
- (1925) *The Great Gatsby*. New York: Scribner.
- Flamini, Roland (1975) *Scarlett, Rhett, and a Cast of Thousands: The Filming of *Gone with the Wind**. New York: Macmillan.
- Foster, Hannah (1986 [1797]) *The Coquette*. New York: Oxford University Press.
- Google Books Ngram Viewer (2011) ngrams.googlelabs.com/info (accessed April 4, 2011).

- Goulart, Ron (1972) *Cheap Thrills: An Informal History of the Pulp Magazines*. New Rochelle, NY: Arlington House.
- Hawthorne, Nathaniel (1962 [1850]) *The Scarlet Letter*. Columbus: Ohio State University Press.
- Hotz, Robert Lee (2010) "World-Wide Web launches: New Google database puts centuries of cultural trends in reach of linguists." *Wall Street Journal*, December 17, online.wsj.com/article/SB10001424052748704073804576023741849922006.html.
- Howe, Daniel Walker (2007) *What Hath God Wrought: The Transformation of America, 1815–1848*. New York: Oxford University Press.
- Hurston, Zora Neale (1937) *Their Eyes Were Watching God*. Philadelphia: Lippincott.
- Jameson, Fredric (1991) *Postmodernism; or, The Cultural Logic of Late Capitalism*. Durham, NC: Duke University Press.
- (1998) *The Cultural Turn: Selected Writings on the Postmodern, 1983–1998*. New York: Verso.
- Jones, Daryl (1978) *The Dime Novel Western*. Bowling Green, KY: Popular Press.
- Kaplan, Carla (2010) "Making it new: Constructions of modernism," in Paul Lauter (ed.) *A Companion to American Literature and Culture*. Malden, MA: Wiley-Blackwell: 40–56.
- Kaplan, Justin (1966) *Mr. Clemens and Mark Twain: A Biography*. New York: Simon and Schuster.
- Kete, Mary Louise (2000) *Sentimental Collaboration: Mourning and Middle-Class Identity in Nineteenth-Century America*. Durham, NC: Duke University Press.
- Krauth, Leland (1999) *Proper Mark Twain*. Athens: University of Georgia Press.
- Mailer, Norman (1948) *The Naked and the Dead*. New York: Rinehart.
- (1979) *The Executioner's Song*. Boston: Little, Brown.
- Marx, Leo (1964) *The Machine in the Garden: Technology and the Pastoral Ideal in America*. New York: Oxford University Press.
- May, Henry F. (1959) *The End of American Innocence: The First Years of Our Own Time, 1912–1917*. New York: Oxford University Press.
- McCarthy, Cormac (2006) *The Road*. New York: Vintage.
- McHale, Brian (1987) *Postmodernist Fiction*. New York: Methuen.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden (2010) "Supporting online material for quantitative analysis of culture using millions of digitized books." *Science*, December.
- (2011) "Quantitative analysis of culture using millions of digitized books." *Science*, January 14, 176–82.
- Mintz, Steven (2004) *Huck's Raft: A History of American Childhood*. Cambridge, MA: Harvard University Press.
- Morrison, Toni (1987) *Beloved*. New York: Knopf.
- Novak, Barbara (2007) *Nature and Culture: American Landscape and Painting, 1825–1875*, 3rd ed. New York: Oxford University Press.

- Parrington, Vernon Louis (1927) *Main Currents in American Thought: An Interpretation of American Literature from the Beginnings to 1920*. New York: Harcourt, Brace.
- Rado, Lisa, ed. (1994) *Rereading Modernism: New Directions in Feminist Criticism*. New York: Garland.
- Reinier, Jacqueline S. (1996) *From Virtue to Character: American Childhood, 1775–1850*. New York: Twayne.
- Reynolds, David S. (1989) *Beneath the American Renaissance: The Subversive Imagination in the Age of Emerson and Melville*. Cambridge, MA: Harvard University Press.
- Ringe, Donald A. (1971) *The Pictorial Mode: Space and Time in the Art of Bryant, Irving, and Cooper*. Lexington: University of Kentucky Press.
- Roth, Philip (1997) *American Pastoral*. Boston: Houghton Mifflin.
- Rowson, Susanna (2011 [1791]) *Charlotte Temple*. New York: Norton.
- Russo, Richard (2001) *Empire Falls*. New York: Vintage.
- Salinger, J. D. (1951) *The Catcher in the Rye*. Boston: Little, Brown.
- Santayana, George (1967 [1911]) “The genteel tradition in American philosophy,” in Douglas L. Wilson (ed.) *The Genteel Tradition: Nine Essays by George Santayana*. Cambridge, MA: Harvard University Press: 37–64.
- Sedgwick, Catharine Maria (1998 [1827]) *Hope Leslie*. New York: Penguin.
- Simpson, John A., ed. (2012) *Oxford English Dictionary*, www.oed.com (accessed April 10, 2012).
- Smith, Henry Nash (1978) *Democracy and the Novel: Popular Resistance to Classic American Writers*. New York: Oxford University Press.
- Steiner, Wendy (1999) “Postmodern fictions, 1970–1990,” in Sacvan Bercovitch (ed.) *The Cambridge History of American Literature*. Vol. 7, *Prose Writing, 1940–1990*. New York: Cambridge University Press: 427–538.
- Susman, Warren I. (1984) *Culture as History: The Transformation of American Society in the Twentieth Century*. New York: Pantheon.
- Tompkins, Jane (1985) *Sensational Designs: The Cultural Work of American Fiction, 1790–1860*. New York: Oxford University Press.
- Twain, Mark (1972 [1876]) *Tom Sawyer*. New York: Pocket.
- (2003 [1884]) *Huckleberry Finn*. Berkeley: University of California Press.
- Welter, Barbara (1966) “The cult of true womanhood, 1820–1860.” *American Quarterly* 18 (2, pt. 1): 151–74.
- Wharton, Edith (1905) *House of Mirth*. New York: Scribner.